

SEQUENCE LISTING

<110> Strathmann AG & Co.

<120> Virus-Vaccine

<130> P057760

<140>

<141>

<150> 199 07 485.2

<151> 1999-02-12

<160> 12

<170> PatentIn Ver. 2.1

<210> 1

<211> 9709

<212> DNA

<213> Human immunodeficiency virus

<400> 1

tggaggcgt aatttggtcc caaaaaagac aagagatcct tgcgttgg atctaccaca 60  
cacaaggcta cttccctgat tggcagaact acacaccagg gccaggatc agatatccac 120  
tgaccttgg atgggtcttc aagttgtac cagttgaacc agagcaagta gaagaggcca 180  
aataaggaga gaagaacacgc ttgttacacc ctatgagcca gcatgggatg gaggaccgg 240  
agggagaagt attagtgtgg aagtttgaca gcctcttagc atttcgtcac atggcccag 300  
agctgcattcc ggagttactac aaagactgtc gacatcgagc tttctacaag ggactttccg 360  
ctggggactt tccagggagg tggccctgg gcgggactgg ggagtggcga gcccctcagat 420  
gctacatata agcagctgct ttttgcctgt actgggtctc tctggttaga ccagatctga 480  
gcctgggagc tctctggcta actaggaaac ccactgtta agcctaata aagcttgccc 540  
tgagtgtca aagtagtgtg tgccctgtt tgggtgtact ctggtaacta gagatccctc 600  
agaccctttt agtcagtgtg gaaaatctct agcagtggcg cccgaacagg gacttggaaag 660  
cgaaagtaaa gccagaggag atctctcgac gcaggactcg gcttgcgtcaa gcgcgcacgg 720  
caagaggcga ggggcggcga ctggtgagta cgccaaaaat ttgacttagc ggaggctaga 780  
aggagagaga tgggtgcgag agcgtcggtt ttaagcccc gagaattaga taaaatggaa 840  
aaaattcggt taaggccagg ggaaaagaaa caatataaac taaaacatat agtatggca 900  
agcaggaggc tagaacgatt cgcagttaat cctggccctt tagagacatc agaaggctgt 960  
agacaaatac tgggacagct acaaccatcc cttcagacag gatcagaaga acttagatca 1020  
ttatataata caatagcagt cctctattgt gtgcataaaa ggatagatgt aaaagacacc 1080  
aagaaagcct tagataagat agaggaagag caaaacaaaa gtaagaaaaa ggcacagcaa 1140  
gcagcagctg acacaggaaa caacagccag gtcagccaaa attaccctat agtgcagaac 1200  
ctccaggggc aaatggtaca tcagggccata tcaccttagaa ctttaatgc atggtaaaa 1260  
gtagtagaag agaaggctt cagccccgaa gtaataccca tggggatgc attatcagaa 1320  
ggagccaccc cacaagattt aaataccatg ctaaacacag tggggggaca tcaagcagcc 1380  
atgc当地gt taaaagagac catcaatggaa gaagctgcag aatgggatag attgcatacc 1440  
gtgc当地gtc ggcctattgc accaggccag atgagagaac caaggggaag tgacatagca 1500  
ggaactacta gtacccttca ggaacaaaata ggatggatga cacataatcc acctatccca 1560  
gtaggagaaa tctataaaag atggataatc ctgggattaa ataaaatagt aagaatgtat 1620  
agccctacca gcattctggc cataagacaa ggacccaaagg aacccttag agactatgt 1680  
gaccgattct ataaaactct aagagccgag caagcttcac aagaggtaaa aaattggatg 1740  
acagaaaacct tggggccca aaatgcgaac ccagattgtt agactatccc aaaagcattg 1800  
ggaccaggag cgacactaga agaaatgtt acagcatgtc agggagtgccc gggaccggc 1860  
cataaagcaa gagtttggc tgaagcaatg agccaaagtaa caaatccagc taccataatg 1920

atacagaaaag gcaatttttag gaaccaaaga aagactgtta agtgtttcaa ttgtggcaaa 1980  
gaagggcaca tagccaaaaa ttgcagggcc ccttagaaaa aggctgttg gaaatgtgga 2040  
aaggaaggac accaaatgaa agattgtact gagagacagg ctaatttttt agggaaagatc 2100  
tggccttccc acaagggaag gccagggaat ttcttcaga gcagaccaga gccaacagcc 2160  
ccaccagaag agagcttcag gtttggggaa gagacaacaa ctccctctca gaagcaggag 2220  
ccgatagaca aggaactgtta tccttagct tccctcagat cactcttgg cagcgacccc 2280  
tcgtcacaat aaagataggg gggcaattaa aggaagctct attagataca ggagcagatg 2340  
atacagtatt agaagaaatg aatttgccag gaagatggaa accaaaaatg atagggggaa 2400  
ttggaggttt tatcaaagta ggacagtatg atcagatact catagaaaatc tgccgacata 2460  
aagctatagg tacagtatta gtaggaccta cacctgtcaa cataattgga agaaatctgt 2520  
tgactcagat tggctgcact ttaaattttc ccattagtcc tattgagact gtaccagtaa 2580  
aattaaagcc aggaatggat ggcccaaag ttaaacaatg gccattgaca gaagaaaaaa 2640  
taaaagcatt agtagaaatt tgtacagaaa tggaaaagga agaaaaattt tcaaaaattt 2700  
ggcctgaaaaa tccatacaat actccagttt ttgcataaaa gaaaaaaagac agtactaaat 2760  
ggagaaaattt agtagatttc agagaactta ataagagaac tcaagatttc tgggaagtcc 2820  
aatttaggaat accacatcct gcagggttaa aacagaaaaa accatgttca gtaactggatg 2880  
tggcgatgc atattttca gttcccttag ataaagactt caggaagtat actgcattta 2940  
ccatacctag tataaacaat gagacaccag ggatttagata tcagtagataat gtgcctccac 3000  
aggatggaa aggatcacca gcaatattcc agtgttagcat gacaaaaatc ttagagcctt 3060  
ttagaaaaca aaatccagac atagtcatct atcaatataat ggtatgattt tatgttaggat 3120  
ctgacttaga aatagggcag catagaacaa aaatagagga actgagacaa catctgttga 3180  
gggggggatt taccacacca gacaaaaaac atcagaaaaga acctccattc ctttgatgg 3240  
gttatgaact ccattctgtt aaatggacag tacagctat agtgcgtgca gaaaaggaca 3300  
gctgactgt caatgacata cagaaattt tagggaaaattt gaattgggca agtcagattt 3360  
atgcagggat taaagttaagg caattatgtt aacttcttag ggaacccaa gcaactaacag 3420  
aagttagtacc actaacagaa gaagcagagc tagaactggc agaaaaacagg gagattctaa 3480  
aagaaccggt acatggatgt tattatgacc catcaaaaga cttaatagca gaaatacaga 3540  
agcaggggca aggccaatgg acatataaaa ttatcaaga gccatttaaa aatctgaaaa 3600  
cagggaaata tgcaagaatg aagggtgccc acactaatga tggaaaacaa ttaacagagg 3660  
cagtacaaaa aatagccaca gaaagcatag taatatgggg aaagactcct aaatttaat 3720  
taccatataa aaagggaaaca tggaaagcat ggtggacaga gtatggca gcccacccgg 3780  
ttcctgagtg ggagtttgc aatccccctc ctttagtgaa gtatggatc cagttagaga 3840  
aagaaccat aataggagca gaaactttct atgttagatgg ggcagccaaat agggaaacta 3900  
aatttagaaaa agcaggatataat gtaactgaca gaggaagaca aaaagttgtc cccctaacgg 3960  
acacaacaaa tcagaagact gagttacaag caattcatct acgtttgcag gattcgggat 4020  
tagaagtaaa catagtgaca gactcacaat atgcatttggg aatcattcaa gcacaaccag 4080  
ataagagtga atcagagttt gtcagtcaaa taatagagca gtaataaaaa aaggaaaaaa 4140  
tctacctggc atgggtacca gcacacaaaag gaattggagg aaatgaacaa gtagatgggt 4200  
tggtcagtgc tggaaatcagg aaagtactat tttagatgg aatagataag gcccagaag 4260  
aacatgagaa atatcacagt aattggagag caatggctag tgattttaac ctaccacctg 4320  
tagtagcaaa agaaatagta gccagctgtg ataaatgtca gctaaaaggg gaagccatgc 4380  
atggacaagt agactgttagc ccaggaatat ggcagctaga ttgtacacat tttagaaggaa 4440  
aagttatctt ggttagcattt catgtgcca gtggatataat agaagcagaa gtaattccag 4500  
cagagacagg gcaagaaaaca gcataacttcc tcttaaaattt agcaggaaga tggccagtaa 4560  
aaacagtaca tacagacaat ggcagcaattt tcaccagttac tacagttaa gccgcctgtt 4620  
ggtggccggg gatcaagcag gaatttggca ttccctacaa tccccaaatg caaggagtaa 4680  
tagaatctat gaataaaagaa ttaaagaaaa ttataggaca ggttaagagat caggctgaac 4740  
atcttaagac agcagtacaa atggcagttt tcattccacaa tttaaaaaga aaagggggga 4800  
ttggggggtt cagtgcaggg gaaagaatag tagacataat agcaacagac atacaaacta 4860  
aagaattaca aaaacaaaattt acaaaaatttcc aaaattttcg gtttattac agggacagca 4920  
gagatccagt ttggaaagga ccagcaaagc tcctctggaa aggtgaaggg gcagtagtaa 4980  
tacaagataa tagtgacata aaagttagtgc caagaagaaa agcaaagatc atcaggattt 5040  
atggaaaaca gatggcaggat gatgattgtg tggcaagtag acaggatgag gattaacaca 5100  
tggaaaagat tagaaaaca ccatatgtat atttcaagga aagctaagga ctggttttat 5160  
agacatcact atgaaaagatc taatccaaaa ataagttcag aagtacacat cccactaggg 5220  
gtgtctaaat tagtaataac aacatattgg ggtctgcata caggagaaag agactggcat 5280  
ttgggtcagg gagtctccat agaatggagg aaaaagagat atagcacacaca agtagaccct 5340

gacctagcag accaactaat tcatctgcac tattttgatt gttttcaga atctgctata 5400  
agaatatacca tatttagcag tatagttgt cctagggttg aatatcaagc aggacataac 5460  
aaggtaggat ctctacagta cttggcacta gcagcattaa taaaacccaa acagataaaag 5520  
ccaccttgc ctatgttag gaaactgaca gaggacagat ggaacaagcc ccagaagacc 5580  
aagggccaca gagggagcca tacaatgaat ggacactaga gcttttagag gaacttaaga 5640  
gtgaagctgt tagacattt cctaggatat ggctccataa cttaggacaa catatctatg 5700  
aaacttacgg ggatacttgg gcaggagtgg aagccataat aagaattctg caacaactgc 5760  
tgtttatcca tttcagaatt ggggtcgac atagcagaat aggcgttact cgacagagga 5820  
gagcaagaaa tggagccagt agatcctaga cttagcccc ggaagcatcc aggaagtca 5880  
cctaaaactg cttgtaccaa ttgttattgt aaaaagtgtt gctttcattg ccaagttgt 5940  
ttcatgacaa aagccttagg catctccat ggcaggaaga agcggagaca ggcacgaa 6000  
gctcatcaga acagtcaagc tcacatcaagct tctctatcaa agcagtaagt agtacatgt 6060  
atgcaaccta taatgttagc aatagtagca ttagtagtag caataataat agcaatagtt 6120  
gtgtggtcca tagtaatcat agaatatagg aaaatattaa gacaaagaaa aatagacagg 6180  
ttaattgata gactaataga aagagcagaa gacagtggca atgagagtga aggagaagta 6240  
tcagcacttg tggagatggg ggtggaaatg gggcaccatg ctccctggg tattgtatg 6300  
ctgttgtgct acagaaaaat tggggtcac agtctattat ggggtacctg tggggagga 6360  
agcaaccacc actcttattt gtgcacatcaga tgctaaagca tatgatacag aggtacataa 6420  
tgttgggcc acacatgcct gtgtacccac agaccccaac ccacaagaag tagtattgt 6480  
aaatgtgaca gaaaattta acatgtggaa aaatgcacatg gtggacacaga tgcacggg 6540  
tataatcagt ttatgggatc aaagcctaaa gccatgtgtaa attaaaccc cactctgtt 6600  
tagtttaaag tgcactgtt tgaagaatga tactaatacc aatagtagta gcgggagaat 6660  
gataatggag aaaggagaga taaaaaactg ctcttcaat atcagcacaa gcataagaga 6720  
taaggtgcag aaagaatatg cattttttt taaactgtat atagtagccaa tagataatac 6780  
cagctatagg ttgataagtt gtaacaccc agtcattaca caggcctgtc caaaggatc 6840  
ctttgagcca attcccatac attattgtgc cccggctggt tttgcattt taaaatgtaa 6900  
taataagacg ttcaatggaa caggaccatg tacaatgtc agcacagttt aatgtacaca 6960  
tggaatcagg ccagtagtat caactcaact gctttttaaat ggcagttctg cagaagaaga 7020  
tgttagtaatt agatctgcca atttcacaga caatgttaaa accataatag tacagctgaa 7080  
cacatctgtt gaaattaatt gtacaagacc caacaacaat acaagaaaaa gtatccgtat 7140  
ccagagggga ccagggagag cattttttt aataggaaaa ataggaaaata tgagacaagc 7200  
acattgtaac attagtagag caaaatggaa tgccacttta aaacagatag ctagcaaatt 7260  
aagagaacaa ttggaaata ataaaacaat aatctttaag caatccctcag gaggggaccc 7320  
agaaattgtt acgcacagtt ttaattgtgg agggaaattt ttctactgtt attcaacaca 7380  
actgtttat agtacttggt ttaatgttac ttggagttactt gaaagggtcaaa ataacactga 7440  
aggaagtgc acaatcacac tcccatgcag aataaaacaa ttataaaaca tgtggcagga 7500  
agttagaaaa gcaatgtatg cccctcccat cagggacaa attagatgtt catcaaataat 7560  
tactgggctg ctattaacaa gagatggtgg taataacaac aatgggtccg agatcttcag 7620  
acctggagga ggcgatatac gggacaattt gagaagtggaa ttatataaaat ataaagtgt 7680  
aaaaatttggaa ccatttaggat tagcacccac caaggccaaag agaagagtgg tgcagagaga 7740  
aaaaagagca gtggaaatag gagttttttt cttgggttc ttggggcagc caggaagcac 7800  
tatggctgc acgtcaatgc cgctgacggg acaggccaga caattttgt ctgatatagt 7860  
gcaggcagcag aacaatttgc tgagggctat tgaggccaa cagcatctgt tgcaactcac 7920  
agtctggggc atcaaacagc tccaggcaag aatccctggct gtggaaagat acctaaagga 7980  
tcaacagctc ctggggattt ggggttgc tggaaaactc atttgcacca ctgctgtgcc 8040  
ttggaaatgtt agttggagta ataaatctt ggaacagatt tggataaca tgacctggat 8100  
ggaggggac agagaaaatc acaatttacac aagcttataa cactccttaa ttgaagaatc 8160  
gcaaaaccag caagaaaaga atgaacaaga attattggaa ttagataaaat gggcaagttt 8220  
gtggaaattgg tttaacataa caaattggct gtggatataa aaattattca taatgtatgt 8280  
aggaggctt gtaggtttaa gaatagttt tgctgtactt tctatagtga atagagttt 8340  
gcaggatat tcaccattat cgtttcagac ccaccccttccaa atcccgaggg gacccgacag 8400  
gccccaaaggaa atagaagaag aaggtggaga gagagacaga gacagatcca ttcgattgt 8460  
gaacggatcc ttagcactt tctgggacga tctgcggacg ctgtgcctt tcagctacca 8520  
ccgcttgaga gacttactt tgattgttac gaggattgtt gaaacttctgg gacgcagggg 8580  
gtggaaagcc ctcaaataattt ggtggaaatct cctacagtt tggagtcaagg aactaaagaa 8640  
tagtgcgtt aacttgcgtca atgccacagc catagcgtt gctgaggggc cagatagggt 8700  
tatagaagta ttacaagcag cttatagac tatttgcac atacctagaa gaataagaca 8760

gggcttggaa aggattttgc tataagatgg gtggcaagt gtcaaaaagt agtgtgattg 8820  
gatggcctgc tctaaggaa agaatgagac gagctgagcc acgacagat ggggtgggag 8880  
cagtatctcg agacctagaa aaacatggag caatcacaag tagcaataca gcagctaaca 8940  
atgcgtctg tgcctggcta gaagcacaag aggaggaaga ggtgggttt ccagtcacac 9000  
ctcaggtacc tttaagacca atgacttaca aggcatgtt agatcttagc cacttttaa 9060  
aagaaaaggg gggactggaa ggcttaattc actcccaaag aagacaagat atccttgatc 9120  
tgtgatcta ccacacacaa ggctacttcc ctgattggca gaactacaca ccagggccag 9180  
gggtcagata tccactgacc tttggatggt gctacaagct agtaccagtt gagccagata 9240  
aggtagaaga ggccaataaaa ggagagaaca ccagcttgtt acaccctgtg agcctgcatg 9300  
gaatggatga ccctgagaga gaagtgttag agtggaggtt tgacagccgc ctagcatttc 9360  
atcacgtggc ccgagagctg catccggagt acttcaagaa ctgctgacat cgagcttgct 9420  
acaagggact ttccgctggg gacttccag ggaggcgtgg cctggggggg actggggagt 9480  
ggcgagccct cagatgctgc atataagcag ctgcttttgc cctgtactgg gtctctctgg 9540  
tttagaccaga tctgagcctg ggagctctggcttaactag ggaacccact gcttaagcct 9600  
caataaagct tgccttgagt gcttcaagta gtgtgtgccc gtctgttggt tgactctgg 9660  
aactagagat ccctcagacc cttttagtca gtgtggaaaaa tctctagca 9709

<210> 2  
<211> 854  
<212> PRT  
<213> Human immunodeficiency virus

<220>  
<223> Envelope Polyprotein

<400> 2  
Met Arg Val Lys Glu Lys Tyr Gln His Leu Trp Arg Trp Gly Trp Lys  
1 5 10 15  
  
Trp Gly Thr Met Leu Leu Gly Ile Leu Met Ile Cys Ser Ala Thr Glu  
20 25 30  
  
Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala  
35 40 45  
  
Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu  
50 55 60  
  
Val His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn  
65 70 75 80  
  
Pro Gln Glu Val Val Leu Val Asn Val Thr Glu Asn Phe Asn Met Trp  
85 90 95  
  
Lys Asn Asp Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp  
100 105 110  
  
Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Ser  
115 120 125  
  
Leu Lys Cys Thr Asp Leu Lys Asn Asp Thr Asn Thr Asn Ser Ser Ser  
130 135 140  
  
Gly Arg Met Ile Met Glu Lys Gly Glu Ile Lys Asn Cys Ser Phe Asn  
145 150 155 160

Ile Ser Thr Ser Ile Arg Asp Lys Val Gln Lys Glu Tyr Ala Phe Phe  
165 170 175

Tyr Lys Leu Asp Ile Val Pro Ile Asp Asn Thr Ser Tyr Arg Leu Ile  
180 185 190

Ser Cys Asn Thr Ser Val Ile Thr Gln Ala Cys Pro Lys Val Ser Phe  
195 200 205

Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu  
210 215 220

Lys Cys Asn Asn Lys Thr Phe Asn Gly Thr Gly Pro Cys Thr Asn Val  
225 230 235 240

Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro Val Val Ser Thr Gln  
245 250 255

Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu Asp Val Val Ile Arg Ser  
260 265 270

Ala Asn Phe Thr Asp Asn Ala Lys Thr Ile Ile Val Gln Leu Asn Thr  
275 280 285

Ser Val Glu Ile Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser  
290 295 300

Ile Arg Ile Gln Arg Gly Pro Gly Arg Ala Phe Val Thr Ile Gly Lys  
305 310 315 320

Ile Gly Asn Met Arg Gln Ala His Cys Asn Ile Ser Arg Ala Lys Trp  
325 330 335

Asn Ala Thr Leu Lys Gln Ile Ala Ser Lys Leu Arg Glu Gln Phe Gly  
340 345 350

Asn Asn Lys Thr Ile Ile Phe Lys Gln Ser Ser Gly Gly Asp Pro Glu  
355 360 365

Ile Val Thr His Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys Asn  
370 375 380

Ser Thr Gln Leu Phe Asn Ser Thr Trp Phe Asn Ser Thr Trp Ser Thr  
385 390 395 400

Glu Gly Ser Asn Asn Thr Glu Gly Ser Asp Thr Ile Thr Leu Pro Cys  
405 410 415

Arg Ile Lys Gln Phe Ile Asn Met Trp Gln Glu Val Gly Lys Ala Met  
420 425 430

Tyr Ala Pro Pro Ile Ser Gly Gln Ile Arg Cys Ser Ser Asn Ile Thr  
435 440 445

Gly Leu Leu Leu Thr Arg Asp Gly Gly Asn Asn Asn Gly Ser Glu

450                    455                    460

Ile Phe Arg Pro Gly Gly Asp Met Arg Asp Asn Trp Arg Ser Glu  
465                    470                    475                    480

Leu Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Val Ala Pro  
485                    490                    495

Thr Lys Ala Lys Arg Arg Val Val Gln Arg Glu Lys Arg Ala Val Gly  
500                    505                    510

Ile Gly Ala Leu Phe Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr Met  
515                    520                    525

Gly Cys Thr Ser Met Thr Leu Thr Val Gln Ala Arg Gln Leu Leu Ser  
530                    535                    540

Asp Ile Val Gln Gln Asn Asn Leu Leu Arg Ala Ile Glu Ala Gln  
545                    550                    555                    560

Gln His Leu Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Ala  
565                    570                    575

Arg Ile Leu Ala Val Glu Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly  
580                    585                    590

Ile Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Ala Val Pro Trp  
595                    600                    605

Asn Ala Ser Trp Ser Asn Lys Ser Leu Glu Gln Ile Trp Asn Asn Met  
610                    615                    620

Thr Trp Met Glu Trp Asp Arg Glu Ile Asn Asn Tyr Thr Ser Leu Ile  
625                    630                    635                    640

His Ser Leu Ile Glu Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu Gln  
645                    650                    655

Glu Leu Leu Glu Leu Asp Lys Trp Ala Ser Leu Trp Asn Trp Phe Asn  
660                    665                    670

Ile Thr Asn Trp Leu Trp Tyr Ile Lys Leu Phe Ile Met Ile Val Gly  
675                    680                    685

Gly Leu Val Gly Leu Arg Ile Val Phe Ala Val Leu Ser Ile Val Asn  
690                    695                    700

Arg Val Arg Gln Gly Tyr Ser Pro Leu Ser Phe Gln Thr His Leu Pro  
705                    710                    715                    720

Ile Pro Arg Gly Pro Asp Arg Pro Glu Gly Ile Glu Glu Glu Gly Gly  
725                    730                    735

Glu Arg Asp Arg Asp Arg Ser Ile Arg Leu Val Asn Gly Ser Leu Ala  
740                    745                    750

Leu Ile Trp Asp Asp Leu Arg Ser Leu Cys Leu Phe Ser Tyr His Arg  
755 760 765

Leu Arg Asp Leu Leu Ile Val Thr Arg Ile Val Glu Leu Leu Gly  
770 775 780

Arg Arg Gly Trp Glu Ala Leu Lys Tyr Trp Trp Asn Leu Leu Gln Tyr  
785 790 795 800

Trp Ser Gln Glu Leu Lys Asn Ser Ala Val Asn Leu Leu Asn Ala Thr  
805 810 815

Ala Ile Ala Val Ala Glu Gly Thr Asp Arg Val Ile Glu Val Leu Gln  
820 825 830

Ala Ala Tyr Arg Ala Ile Arg His Ile Pro Arg Arg Ile Arg Gln Gly  
835 840 845

Leu Glu Arg Ile Leu Leu  
850

<210> 3

<211> 107

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of the artificial sequence:  
oligonucleotide for cloning

<400> 3

aagatgtagt aatttagatct gccaaattca cagacaatgc taaaaccata atagtgacgc 60  
tgaacacatc gttagaaatt aattgtacaa gacccaaacaa caataca 107

<210> 4

<211> 120

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of the artificial sequence:  
oligonucleotide for cloning

<220>

<221> misc\_feature

<222> (97)..(99)

<223> Sequence at this position: (GA)(AT)(GATC), ie.  
base at position 97 can be G or A, base at  
position 98 can be A or T, and base at  
position 99 can be G, A, T or C.

<400> 4

ttttgctcta gaaatgttac aatgtgcttg tcttatgtct cctgttgcag cttctgttgc 60  
atgaaatgct ctccctggtc cgatatggat actatgrwnt tttcttgtat tggtgttggg 120

<210> 5  
<211> 17  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of the artificial sequence:  
sequencing primer

<400> 5  
ccatgtacaa atgtcag 17

<210> 6  
<211> 17  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of the artificial sequence:  
sequencing primer

<400> 6  
aaaactgtgc gttacaa 17

<210> 7  
<211> 17  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of the artificial sequence:  
sequencing primer

<400> 7  
gtaaaaacgac ggccagt 17

<210> 8  
<211> 17  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of the artificial sequence:  
sequencing primer

<400> 8  
cagggaaacag ctatgac 17

<210> 9  
<211> 2148

<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of the artificial sequence: synthetic DNA

<220>  
<221> misc\_feature  
<222> (3)..(9)  
<223> BstEII cleavage site

<220>  
<221> misc\_feature  
<222> (2143)..(2148)  
<223> BamHI cleavage site

<400> 9

tgggtcaccg tctattatgg ggtgcctgtg tggaaggaaag caaccaccac tctattttgt 60  
gcatcagatg ctaaagcata tgatacagag gtacataatg ttggggccac acatgcctgt 120  
gtaccacag accccaaccc acaagaagta gtattggtaa atgtgacaga aaattttaac 180  
atgtggaaaa atgacatggt agaacagatg catgaggata taatcagttt atgggatcaa 240  
agcctaagc catgtgtaaa attaaccacca ctctgtgtt gttaaagtg cactgattt 300  
aagaatgata ctaataccaa tagtagtagc gggagaatga taatggagaa aggagagata 360  
aaaaactgca gcttcatat cagcacaagc ataagagata aggtgcagaa agaatatgca 420  
ttctttata aacttgatat agtaccaata gataatacca gctataggtt gataagggt 480  
aacacctcag tgatcacaca ggcctgtcca aaggtatcc ttgagccaaat tcccatacat 540  
tattgtgccc cggctggttt tgcgattcta aaatgtataataaagacgtt caatggaaaca 600  
ggaccatgta caaatgtcag cacagtacaa tgtacacatg gaattcgacc agtagtatca 660  
actcaactgc tggtaaatgg cagtcgtca gaagaagatg tagtaatttag atctgc当地 720  
ttcacagaca atgctaaaac cataatagta cagctgaaca catctgtaga aattaattgt 780  
acaagaccca acaacaatac aagaaaaagt atccgtatcc agaggggacc agggagagca 840  
tttgttacaa tagaaaaat agggaaatatg agacaagcac attgttaacat ttctagagca 900  
aaatggaatg ccactttaaa acagatagct agcaaaattaa gagaacaattt tggaaataat 960  
aaaacaataa tctttaagca gtcatccgga ggggaccacag aaattgtataac gcacagttt 1020  
aattgtggag gggaaattttt ctactgtaat tcaacacaac tggtaatag tacttggtt 1080  
aatagtactt ggagtactga agggtcaaattt aacactgaag gaagtgcac aatcacactc 1140  
ccatgcagaa taaaacaatttataaacatg tggcaggaag tagaaaaagc aatgtatgcc 1200  
cctcccatca gtggccaaat tagatgttca tcaaataattt ctggcgtgtt attaactcga 1260  
gatggtggtt ataacaacaa tgggtccgag attttcagac ctggaggagg cgatatgagg 1320  
gataatttga gaagtgaattt atataaaat aaagttagaa aaattgtaaacc attaggagta 1380  
gcaccacca aggcaaaagag acgcgtggtg cagagagaaa agcgcgcagt gggatagga 1440  
gctctgttcc ttgggttctt gggagcagca ggaagcacta tggcgcgcgt gtcaatgacg 1500  
ctgacgggtac aggccagaca attattgtct gatatagtgc agcagcagaa caatttgctg 1560  
agggcaatttgg aggcgcacaaca gcatctgttcaactcacag tctggggcat caaacagctc 1620  
caggcaagaa tccctggctgtt gggaaagatac ctaaaggatc aacagcttcc ggggatttgg 1680  
ggttgtctgtt gaaaactcat ttgcaccact gctgtgcctt ggaatgttag ttggagtaat 1740  
aaatctctgg aacagatttgg gaataacatg acctggatgg agtggggacag agaaattaac 1800  
aattacacaa gcttaataca ctccttaattt gaagaatcgc aaaaccagca agaaaagaat 1860  
gaacaagaat tatttggattt agataaaatgg gcaagttgtt ggaatttggtaat 1920  
aatttggctgtt ggtatataaa attattcata atgatagttag gaggcttggt aggttttaaga 1980  
atagtttttgc ctgtactttc tataatgtaaat agagtttaggc agggatattc accattatcg 2040  
tttcagaccc acctcccaat cccgagggga cccgacagggc ccgaaggaaat agaagaagaa 2100  
ggtggagaga gagacagaga cagatccattt cgattgtga acggatcc 2148

<210> 10  
<211> 6229

<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of the artificial sequence: synthetic DNA

<220>  
<221> sig\_peptide  
<222> (1293)..(1295)  
<223> env ATG

<220>  
<221> misc\_feature  
<222> (1377)..(1379)  
<223> env AGT, gp120 start

<220>  
<221> misc\_feature  
<222> (1397)..(1403)  
<223> BstEII cleavage site

<220>  
<221> misc\_feature  
<222> (3537)..(3542)  
<223> BamHI cleavage site

<220>  
<221> misc\_feature  
<222> (3855)..(3857)  
<223> env TAA, stop

<400> 10  
ctgacgcgcc ctgttagcggc gcattaagcg cggcgggtgt ggtggttacg cgcaagggtga 60  
ccgcctacact tgccagcggcc cttagcgcccc ctcccttcgc ttcttccct tccttctcg 120  
ccacgttcgc cggctttccc cgtcaagctc taaatcgggg gctcccttta gggttccgat 180  
ttagtgcctt acggcacctc gacccaaaaa aacttgatta gggtgatggg tcacgttagt 240  
ggccatcgcc ctgatagacg gttttcgcc ctttgacggtt ggagtccacg ttcttaata 300  
gtggactctt gttccaaact ggaacaacac tcaaccctat ctcggtctat tctttgatt 360  
tataaggat tttgccgatt tcggcctatt ggttaaaaaa tgagctgatt taacaaaaat 420  
ttaacgcgaa ttttaacaaa atattaacgc ttacaatttc cattcgccat tcaggctgcg 480  
caactgttgg gaagggcgat cggtgccggc ctcttcgcta ttacgcccacg tggcgaaagg 540  
gggatgtgct gcaaggcgat taagttgggt aacgcccaggg tttcccgat cacgacgtt 600  
taaaacgacg gccagtggac gtctagttat taatagtaat caattacgg gtcattagtt 660  
catagcccat atatggagtt ccgcgttaca taacttacgg taaatggccc gcctggctga 720  
ccgccccaaacg accccccggcc attgacgtca ataatgacgt atgttcccat agtaacgcca 780  
ataggggactt tccattgacg tcaatgggtg gagtatttac ggttaaactgc ccacttggca 840  
gtacatcaag tgtatcatat gccaaggtacg cccccctattt acgtcaatga cggtaaatgg 900  
cccgctggc attatgcccc gtacatgacc ttatggact ttctctacttg gcagtgacatc 960  
tacgtatttag tcattcgctat taccatgggtg atgcgggtt ggcagtgatcat caatggcg 1020  
ggatagcggt ttgactcact gggatttcca agtctccacc ccattgacgt caatggaggt 1080  
ttgtttggc accaaaaatca acgggacttt ccaaaatgtc gtaacaactc cgccccattt 1140  
acgcaaattgg gcggttaggcg tgcgttgg gaggctata taagcagacg tcgttttagt 1200  
aacccgtcaga tcgcctggag acggcatcca cgcgtttt acctccatag aagacaccgg 1260  
gacaattcga gtcgggtacc gtcgacgcca ccatgagagt gaaggagaag tatcagcact 1320  
tgtggagatg ggggtggaaa tggggcacca tgctccctgg gatattgatg atctgttagt 1380  
ctacagaaaa attgtgggtc accgtctatt atggggtacc tgtgtggaaag gaagcaacca 1440  
ccactctatt ttgtgcatca gatgctaaag catatgatac agaggtacat aatgtttggg 1500

ccacacatgc ctgtgtaccc acagacccca acccacaaga agtagtattg gtaaatgtga 1560  
cagaaaattt taacatgtgg aaaaatgaca tggtagaaaca gatgcatacg gataataatca 1620  
gttatggga tcaaaggcta aagccatgtg taaaattaac cccactctgt gttagttaa 1680  
agtgcactga tttgaagaat gatactaata ccaatagtag tagcgggaga atgataatgg 1740  
agaaaggaga gataaaaaac tgctcttca atatcagcac aagcataaga gataaggtgc 1800  
agaaagaata tgcattctt tataaacttg atatagtacc aatagataat accagctata 1860  
ggttgataag ttgtaacacc tcagtcatta cacaggccctg tccaaaggtt tccttgagc 1920  
caattcccat acattattgt gccccggctg gtttgcgt tctaaaatgt aataataaga 1980  
cgttcaatgg aacaggacca tgacaaaatg tcagcacagt acaatgtaca catggaatca 2040  
ggccagtagt atcaactcaa ctgctgtt aaaccataat agtacagctg aacacatctg 2100  
ttagatctgc caatttcaca gacaatgcta aaaccataat agtacagctg aacacatctg 2160  
tagaaattaa ttgtacaaga cccacaaca atacaagaaa aagtatccgt atccagaggg 2220  
gaccagggag agcatttgtt acaataggaa aaatagaaaa tatgagacaa gcacattgt 2280  
acattagtag agcaaaaatgg aatgccactt taaaacagat agtacagctg aataagagaac 2340  
aatttggaaa taataaaaaca ataatcttta agcaatccc aggagggggac ccagaaattg 2400  
taacgcacag ttttaattgt ggaggggaat tttctactg taattcaaca caactgttta 2460  
atagacttg gtttaatagt acttggagta ctgaagggtc aaataacact gaaggaagtg 2520  
acacaatcac actcccatgc agaataaaaac aatttataaa catgtggcag gaagtaggaa 2580  
aagcaatgta tgcccccccc atcagtggtc aaatttagatg ttcatcaaat attactggc 2640  
tgctttaaac aagagatggt ggtataaca acaatgggtc cgagatctt agacctggag 2700  
gagggcataat gagggacaaat tggagaagtg aattatataa atataaaagta gtaaaaattg 2760  
aaccattagg agtagcaccc accaaggcaa agagaagagt ggtgcagaga gaaaaaagag 2820  
cagtggaat aggagcttt ttccttgggt tcttggggc agcaggaagc actatggc 2880  
gcacgtcaat gacgctgacg gtacaggcca gacaattttt gtcgtatata gtgcagcagc 2940  
agaacaattt gctgagggctt attgaggcgc aacagcatctt gtgcacactc acagttctggg 3000  
gcatcaaaca gctccaggca agaattcttggc ttgtggaaag atacctaaag gatcaacagc 3060  
tcctggggat ttgggggtgc tctggaaaac tcatttgcac cactgtgtg ccttggaaatg 3120  
ctagttggag taataaatct ctgaaacaga tttggataa catgacctgg atggagttggg 3180  
acagagaaat taacaattac acaagcttaa tacactccctt aatttgaagaa tcgcaaaacc 3240  
agcaagaaaaa gaatgaacaa gaattattgg aatttagataa atgggcaagt ttgtgaaatt 3300  
ggttaacat aacaaatgg ctgtggtata taaaattttt cataatgata gtaggaggct 3360  
ttgttaggtt aagaatagtt ttgtgttac ttctatagt gaatagagtt aggccaggat 3420  
attcaccatt atcgtttcag acccaccccttcc caatcccgag gggacccgac aggcccgaaag 3480  
gaatagaaga agaagggtgga gagagagaca gagacagatc cattcgatata gtgaacggat 3540  
cctttagcact tatctgggac gatctgcgga gcctgtgttctt ctcagctac caccgttgc 3600  
gagacttact ttgtattgtt acgaggattt tggaaacctt gggacgcagg ggggtggaaag 3660  
ccctcaaata ttgggtggaaat ttgttacagt attggagtca ggaactaaag aatagtgttgc 3720  
ttaacttgc tcaatgccaca gccatagcag tagctgaggg gacagatagg gttatagaag 3780  
tattacaagc agcttataaga gctattcgcc acataacctt aagaataaga cagggcttgg 3840  
aaaggatttt gctataagat gggtggcaag tggtaaaaaaa gtagtgtgtat gggatggcct 3900  
gctgttaaggaa aaagaatgg acgagctgag ccagcagcag atgggggtggg agcagttatct 3960  
cgagatctttagtacttacta gcttcgtatcc agacatgata agatacatgg atgagtttgg 4020  
acaaaaccaca actagaatgc agtggaaaaaa atgctttttt tttgtttttt gttgtatctat 4080  
tgcttttattt gtaaccatataa taagctgcaaa taaacaaggat aacaacaaca attgcattca 4140  
ttttatgtttt caggttcagg gggaggtgtg ggaggtttttt taaagcaagt aaaacctcta 4200  
caaatgtgtt atggctgattt atgatccctgc ctcgcgcgtt tcgggtatgc cggtaaaaac 4260  
ctctgacaca tgcagctccc ggagacggc acagctgttgc tggtaagcgga tgccgggagc 4320  
agacaagccc gtcaggccgc gtcagcggtt gttggcggtt gtcggggcgc agccatgacc 4380  
cagtcacgtt gcgatagcgg agtgtataactt ggcttaacta tgcggcatca gagcagattt 4440  
tactgagagt gcaccatata gtcggccgcg ttgtgtgtt gtttccatag gctccggccc 4500  
cctgacgagc atcacaaaaaa tgcacgctca agtcaagaggtt ggcgaaaccc gacaggacta 4560  
taaagataacc aggcgttcc cccttggaaatc tccctgtgc gctctcctgt tccgaccctg 4620  
ccgcattaccg gataacctgtc cgcccttctc ctttcgggaa gctgtggcgtt ttctcatagc 4680  
tcacgctgtt ggtatctcag ttccgggtttag gtcgttgcgtt ccaagctggg ctgtgtgcac 4740  
gaaccccccgg ttcagcccgaa ccgcgtgcgc ttatccggta actatcgtct tgagttcaac 4800  
ccggtaagac acgacttatac gccactggca gcagccactt gtaacaggat tagcagagcg 4860  
aggtatgttag gcggtgttac agagtttttgc aagtggtggc ctaactacgg ctacactaga 4920

aggacagtat ttgttatctg cgctctgctg aagccagttt cttcgaaaa aagagtgg 4980  
agctcttgat ccggcaaca aaccaccgct ggtagcggtg gttttttgt ttgcaagcag 5040  
cagattacgc gcagaaaaaa aggatctcaa gaagatcctt tgatctttc tacgggtct 5100  
gacgctcagt ggaacgaaaa ctcacgttaa gggattttgg tcatgagatt atcaaaaagg 5160  
atcttcacct agatccttt aaattaaaaa tgaagttta aatcaatcta aagtatatat 5220  
gagtaaactt ggtctgacag ttaccaatgc ttaatcagtg aggcacctat ctcagcgatc 5280  
tgtctatitc gttcatccat agttgcctga ctccccgtcg tgtagataac tacgataacgg 5340  
gagggcttac catctggccc cagtgctgca atgataccgc gagacccacg ctcaccggct 5400  
ccagatttt cagcaataaa ccagccagcc ggaaggggcg agcgcagaag tggccctgca 5460  
actttatccg cctccatcca gtctattaat tggccggg aagcttagagt aagttagttcg 5520  
ccagttataa gtttgcgcaa cggtgttgc attgctacag gcatcgtggt gtcacgctcg 5580  
tcgttggta tggcttcatt cagctccggt tcccaacgat caaggcgcagt tacatgatcc 5640  
cccatgttgc gcaaaaaaagc ggttagctcc ttccgtcctc cgatcgttgc cagaagtaag 5700  
ttggccgcag tgttatcaact catggttatg gcagcactgc ataattctct tactgtccatg 5760  
ccatccgtaa gatgctttc tgtgactggg gagtactcaa ccaagtcatt ctgagaatag 5820  
tgtatgcggc gaccgagttt ctcttgcggc gctcaatac gggataatac cgcgccacat 5880  
agcagaactt taaaagtgc catcatttggaa aacgttctt cggggcgaaa actctcaagg 5940  
atcttaccgc tggtagatc cagttcgatg taacccactc gtgcacccaa ctgatcttca 6000  
gcatcttttta ctttccaccag cgtttctggg tgagcaaaaaa caggaaggca aaatgcccga 6060  
aaaaaggaa taagggcgac acggaaatgt tgaataactca tacttttctt tttcaatat 6120  
tattgaagca ttatcaggg ttattgtctc atgagcggat acatatttga atgtattttag 6180  
aaaaataaaac aaatagggtt tccgcgcaca ttccccgaa aagtgcac 6229

<210> 11  
<211> 860  
<212> DNA  
<213> Human immunodeficiency virus

<220>  
<221> misc\_feature  
<222> (1)..(860)  
<223> PI-932 original sequence V1-V2-V3-loop

<400> 11  
tgtgtaccca cagaccccaa cccacaaaaa gtagtattgg aaaatgtgac agaaaatttt 60  
aacatgtgga aaaatgacat ggtagaacag atgcattgagg atataatcaa tttatggat 120  
caaaggctaa agccatgtgt aaaactaacc ccactctgtg ttactttaaa ttgcactgat 180  
gctgatttaa attgcaataa tactgattt aattgcacta aagctaattt gggaaaaat 240  
actcataaca atactattag tggaaaata atagagaaaag tagaaataaa aaactgctct 300  
ttcaaggtca ccacaggcat aaggataag atgcaaaaaag aatatgcact tttgaataaa 360  
cttgatatacg taccaataga taatgataag aataataacta actttatatt gataagggt 420  
aacacctcga ccattacaca ggcctgtcca aaggtatcct ttgagccaat tccatatacat 480  
ttttgtgccc cggctggttt tgcgattcta aagtgtaatg aaaagagttt cagtgaaaa 540  
ggaccatgtt aaaaatgtcag cacagtacaa tgtacacatg gaatttaggc agtagtgtca 600  
actcaactgc tggtaatgg cagtcttagca gaaaaagaag tagtaattag atctgagaat 660  
ttcacagaca atgctaaaac cataatagta cagctgaagg aatctgtaaa cattactgt 720  
ataagacccc acaacactgt aacagacagg atacatatacg ggccaggag atcatttcat 780  
acaacaagaa aaataaaagg agatataaga caagcacatt ttagccttag gagaaaaagat 840  
tggataaca ctttacaaga 860

<210> 12  
<211> 870  
<212> DNA  
<213> Artificial Sequence

<220>

<223> Description of the artificial sequence: PI-932  
gene cassette, comprising the cleavage sites for  
restriction enzymes BspT1, PstI, BclI, EcoRI,  
BglII, PvuII, XbaII, NheI

<400> 12

tgtgtaccca cagaccccaa cccacaaaag gtagtattgg aaaatgtgac agaaaatttt 60  
aacatgtgga aaaatgacat ggtagaacag atgcattagg atataatcaa tttatggat 120  
caaagcccta agccatgtgt aaaactaacc ccactctgtt tactttaaa ttgcactgat 180  
gctgatttaa attgcaataa tactgattta aattgcacta aagctaattt gggaaaaat 240  
actcataact gcagtattag tggaaaata atagagaaag tagaaataaa aaactgctct 300  
ttcaaggcata ccacaggcat aaggataag atgcaaaaag aatatgcact tttgaataaa 360  
cttgatatacg taccaataga taatgataag aataatacta accttatatt gataagttgt 420  
aacaccccg tgatcacaca ggctgtcca aaggtatcct ttgagccat tcccatacat 480  
tttgtgccc cggctggttt tgcattcta aagtgtaatg aaaagaggtt cagtggaaaa 540  
ggaccatgtt aaaaatgtcag cacagtacaa tgtacacatg gaattcggcc agtagtgtca 600  
actcaactgc tggtaatgg cagtcttagca gaaaaagaag tagtaattag atctgagaat 660  
ttcacagaca atgctaaaac cataatagta cagctgaagg aatctgtaaa cattacttgt 720  
ataagacccc acaacactgt aacagacagg atacatatacg gcccaggag atcatttcatt 780  
acaacaagaa aaataaaagg agatataaga caagcacatt gtgcctttc tagaaaagat 840  
tggataaca cttaacaaga gatagcttagc 870